**AFRL-RI-RS-TM-2008-1**
**Final Technical Memorandum**
**January 2008**


# 3D SPATIAL MODELING PLAN FOR BIOSPICE

**University of North Carolina - Chapel Hill**



**AIR FORCE RESEARCH LABORATORY**
**INFORMATION DIRECTORATE**
**ROME RESEARCH SITE**
**ROME, NEW YORK**

# NOTICE AND SIGNATURE PAGE



AFRL-RI-RS-TM-2008-1 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

| | |
|---|---|
| /s/ | /s/ |
| ROBERT L. KAMINSKI<br>Work Unit Manager | WARREN H. DEBANY, Jr.<br>Technical Advisor, Information Grid Division<br>Information Directorate |

# REPORT DOCUMENTATION PAGE

*Form Approved*
**OMB No. 0704-0188**



| 1. REPORT DATE *(DD-MM-YYYY)* | 2. REPORT TYPE | 3. DATES COVERED *(From - To)* |
|---|---|---|
| JAN 2008 | Final | Mar 05 – Jun 07 |

**4. TITLE AND SUBTITLE**

3D SPATIAL MODELING FOR BIOSPICE

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
FA8750-05-1-0118

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**

David Adalsteinsson and Phil Colella

**5d. PROJECT NUMBER**
BIOC

**5e. TASK NUMBER**
CH

**5f. WORK UNIT NUMBER**
AP

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
University of North Carolina – Chapel Hill
104 Airport Drive, Ste 2200
Chapel Hill NC 27599-1550

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

AFRL/RIG
525 Brooks Rd
Rome NY 13441-4505

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSORING/MONITORING AGENCY REPORT NUMBER**
AFRL-RI-RS-TM-2008-1

**12. DISTRIBUTION AVAILABILITY STATEMENT**
*APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. PA# WPAFB 08-0059*

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**
This effort developed the spatial modeling component of the Defense Advanced Research Projects Agency (DARPA) BioSpice project. The principal accomplishments include the development of a new class of methods for simulating reaction-diffusion processes in cells, and of an end-to-end methodology for obtaining discretization data from image data via level sets. These methods were tested on several model problems in systems biology.

**15. SUBJECT TERMS**
Bio Computing, Spatial Modeling Partial Differential Equations

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Robert L. Kaminski |
| U | U | U | UL | 10 | 19b. TELEPHONE NUMBER (*Include area code*) N/A |

# Final Report for the 3D Spatial Modeling Plan for the DARPA BioSPICE Project, Grant No. FA8750-05-1-0118 (Lawrence Berkeley National Laboratory, University of North Carolina)

## December 3, 2007

## Summary

We describe here the work done under the spatial modeling component of the DARPA BioSpice project. The principal accomplishments include the development of a new class of methods for simulating reaction-diffusion processes in cells, and of an end-to-end methodology for obtaining discretization data from image data via level sets. These methods were tested on several model problems in systems biology.

## Introduction

Partial differential equations (PDE) based spatial modeling in systems biology offer a number of distinctive difficulties relative to modeling of human-engineered systems. The complexity of the system, both in physical space and in state space is much greater. Furthermore, to use such models requires a high degree of interaction with chemical and image data. Finally, models must have a high degree of ease of use, in order to have any hope of matching the throughput of current techniques for performing laboratory experiments in systems biology. In particular, the traditional approaches to spatial discretization used in engineered systems, such as multiblock and unstructured grids, require a large amount of human (days to weeks) to generate grids that will yield acceptably accurate solutions for such complex systems. Such an approach represents a major obstacle to progress in this field.

In the last decade, there has been substantial progress in a different approach to solving PDE in complex geometries, called variously the cut-cell, Cartesian grid, or embedded boundary approach. In this approach, the grid generation problem is reduced to computing the intersection of the boundary with rectangular grid cells. When applied to the computation of flow around complex aerodynamic shapes, this leads to a grid generation method that takes a few minutes on a high-end workstation (Aftosmis, Berger, and Melton, AIAA J. Jan. 1998). At the same time, there has been the development of new methods for segmentation of image data to obtain a representation of features in that data as the level set of some function (Malladi, Sethian, and Vermuri, IEEE Transactions on Pattern Anal. Machine Intell., 1995).

The goal of this work has been to combine these two techniques to address PDE-based spatial modeling in systems biology. This work had three components.

1. Simulation of reaction-diffusion problems in the cytoplasm and on the membrane using the embedded boundary approach.
2. Embedded boundary grid generation from image data of cells using a level-set representation derived from a front propagation method for segmentation.
3. Coupling of the spatial modeling tools to the BioSpice framework to provide access to chemical rate data.

# Methods, Assumptions and Procedures

## *Simulating Diffusion Processes in Complex Geometries*

Our fundamental approach to spatial discretization is to use finite-volume discretizations on rectangular grids. For parabolic problems such as arise in the cell biology models here, conservation is desirable to maintain accuracy in marginally-resolved settings, particularly in conjunction with chemical reactions. In addition, the natural solvability conditions play a role in recovering the numerical analogue of the thin-layer asymptotics required for our approach to surface diffusion. We use an embedded boundary approach to represent complex geometries with finite-volume discretizations. In this approach, the geometry of the boundary is represented on the grid by intersecting each rectangular Cartesian cell with the irregular boundary. The specific approach taken here is based on a new formal truncation error analysis in which the discretized solution is centered on the rectangular Cartesian mesh, while the various operator discretizations are centered on the appropriate centroids of the intersection of the cell with the domain. The truncation-error analysis is combined with a heuristic modified-equation analysis for singular sources to provide a useful guide for predicting the accuracy of embedded boundary methods. These spatial discretization methods can be used in conjunction with implicit time discretizations, with the resulting large linear systems solved using multigrid iteration. This leads to methods that are both accurate and efficient (both in terms of CPU time and memory).

An important component of the embedded boundary approach is the construction the discretization information ("grid generation") from an implicit function representation of the boundary. In this approach, the surface is represented as the set of all point in space where a scalar function takes on some specified value (e.g. zero). The intersection of each of the rectangular finite volume cells with the volume enclosed by that surface, as well as the various areas and moments on the boundary of the intersection, can be computed using repeated applications of the divergence theorem combined with a least-squares procedure, reducing the problem ultimately to finding the intersection of the surface with the Cartesian coordinate lines. As described below, such implicit function representations can be computed efficiently from data obtained from various types of imaging technologies, such as magnetic resonance imaging, computer tomography, and convolution microscopy.

## *Construction of Implicit Function Representations from Image Data*

Our approach to the construction of implicit function representations from image data is based on the work of Malladi, Sethian, and Vermuri for computing a surface representation from a gray-intensity field. The approach works both in two and three dimensions, and the three-dimensional formulation is a mathematical extension of the two-dimensional formulation. The approach has the following steps.

1. The method starts with an initial shape. In this implementation, the initial shape starts as being strictly larger than the final shape will be. The shape could be the outer box/cube of the image, but is typically picked to be closer to the final shape.

2. The shape is contracted, almost like an elastic membrane. The evolution is implemented using the level set method, which is a front propagation scheme that deals naturally with evolution laws based on partial differential equations. The speed of the membrane motion depends on geometric properties of the membrane, as well as the intensity field given by the image stack. Specifically, the velocity of the front consists of three components: (1) motion in the direction of the normal to the image gradient, (2) an edge detection term, proportional to the gradient of the magnitude of the gradient of the intensity; and (3) a curvature term, to smooth out small-scale variations arising from localized gaps in the data.

3. The shape will reach a steady state when curvature and image-based information counters the overall contraction imposed. The shape will move slowly as this shape is reached. Once this implicit function representation is computed, one can use an eikonal solver to construct a signed distance function, if needed.

## *Software Development*

The starting point for the development of the software for diffusion modeling was LBNL's Chombo package, an object-oriented framework written in C++ to support high-performance parallel implementations of finite-volume methods for partial differential equations. The embedded boundary algorithms are organized as combinations of operations on unions of rectangular arrays, with irregular calculations near the embedded boundary performed on a set of co-dimension one. Chombo is a hybrid programming system that reflects this organization. Higher-level irregular data and control structures and the irregular co-dimension one calculations are implemented using C++, with operations on rectangular arrays implemented in Fortran in order to obtain high performance. The use of Chombo as a primary development platform enabled us to leverage the extensive investment in Chombo by the Department of Energy and NASA.

The code for computing implicit function representations from image data consists of foundation classes and level set classes. The foundation classes are from http://www.visualdatatools.com/DTSource.html, a C++ source class library that can be downloaded and intended for use in numerical programs. The DTSource library does not implement the level set method, but for that a number of functions and classes were implemented, and are here called the "DTExt" library. This extension library implements full and banded level set methods, and uses the DTSource library as the foundation classes as well as to handle input and output. Our approach to the user interface is to use matlab as the cross platform engine to drive the simulations. Matlab has an extensive set of functions built in that help with reading tiff images and display two and three dimensional results. For the matlab implementation, a number of matlab scripts are included that can be used to step through the process, and those scripts can be used for further automation, once parameters have been established. The parameters will not be universal across different image types, but images gotten with the same type of equipment and experimental configuration will require little or no tweaking.
Most of the matlab scripts provided are just wrappers that call the command line utilities. Those have to be compiled separately using the make files provided. It is certainly also possible to merge the codes into a single executable, since all of the routines link against the DTSource library and the DTExt library.

## Results and Discussion

We attempted to integrate into the BioSpice Dashboard the capability to run simulations of phenomena modeled by systems of reaction-diffusion equations. In particular, we focused our efforts on simulations of chemotaxis, or locomotion in response to a chemical gradient, sporulation, the response of cells to stress, and the motion of lipid rafts, a surface phenomena involving diffusion transport and chemical reactions.

This required the development of algorithms and implementations for

1. Representation of diffusive transport in bulk.

2. Representation of diffusive transport in time dependent domains.

3. Representation of diffusive transport on surfaces.

4. Coupling chemical reactions to transport.

This work has resulted in two publications in archival journals of our results on diffusive transport. Furthermore, work begun in this project related to the representation of geometry will soon be submitted for publication.

In addition to core research in numerical algorithms for the solution of elliptic and parabolic differential equations, we committed the necessary software development

resources to make these results accessible to systems biologists using the BioSpice Dashboard, culminating in several live demonstrations at BioComp PI meetings.

## *Representation of Diffusive Transport in Bulk*

Our motivation for studying diffusive transport was problems involving cell signaling, sporualation, and the formation of lipid rafts. For example, Meinhart's often used model of chemotaxis requires local activation on the cell membrane combined with global diffusion of an inhibitor to motion. We developed and implemented an algorithm for diffusive transport in complex geometries, that permitted the testing of this model in a simulation that used realistic representation of the cell, which in our case was a hl-60 cell resembling a neutrophil in the relevant respects. The results from this experiment demonstrated that the model for chemotaxis was less successful in a complex (neutrophil) geometry, and even less successful in three dimensions than it was in a two-dimensional square domain.

## *Representation of Diffusive Transport in Time-Dependent Domains*

We extended this work to the calculation of diffusion on a moving cylinder with spherical caps. This represented a necessary piece of the mathematics behind an abstraction of the sporulation process in Caulobacter. Using a space-time divergence theorem we represented diffusion on a time-dependent domain as a sequence of fixed boundary problems, which we computed using the embedded boundary algorithm for fixed domains.

## *Representation of Diffusive Transport on Surfaces*

The fixed boundary algorithm for diffusion in complex domains proved useful again when we studied problems that involved surface diffusion. For example, many cell signaling processes begin with the diffusion of a species on the cell membrane. We combined our embedded boundary and image processing methodologies to develop a new algorithm for representing diffusion on a surface (Schwartz, et. al., Proceedings of the U.S. National Academy of Sciences, Vol. 102 (2005), pp. 11151-11156). In this approach, diffusion on the surface is replaced by ordinary bulk diffusion, but in an annular volume consisting of all of the points a fixed distance from the surface. Thus the computational domain is specified in terms of an implicit function, and the tools we have developed for grid generation can be applied. In addition, such an implicit function can be computed from image data using the techniques described below. Finally, thin-layer asymptotics and modified equation analysis was used to show that, if the width of the annular region is a constant multiple of the mesh spacing, then as the mesh spacing goes to zero, the method is second-order accurate, a result that was verified by numerical tests.

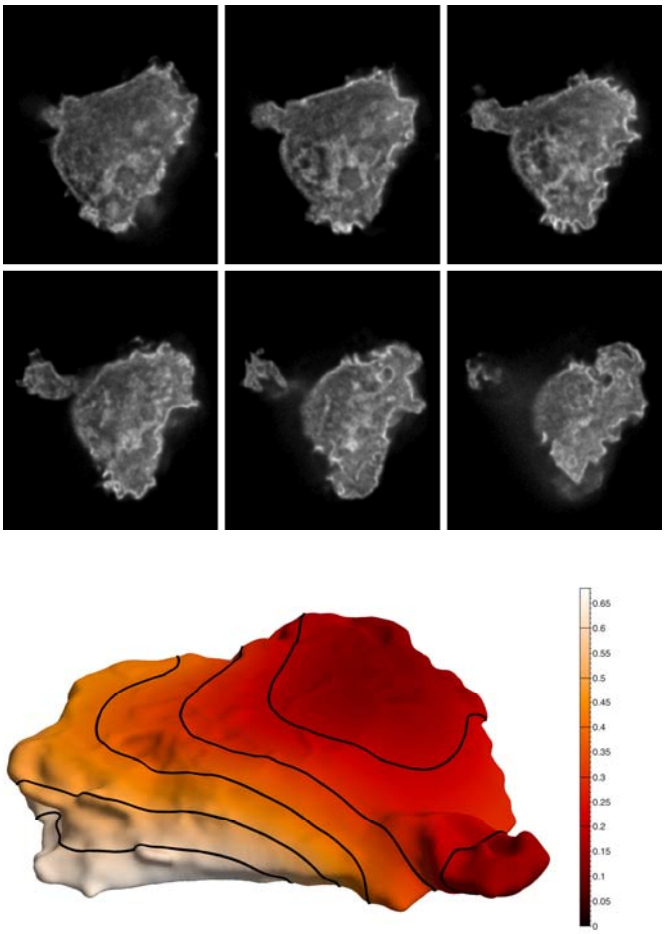


Figure 1. Numerical simulation of surface diffusion on the membrane of an hl-60 cell. Top: convolution-microscopy image data from which the signed-distance function was generated. Bottom: numerical simulation using embedded boundary calculation in an annular region defined by the signed-distance function.

## *Coupling Chemical Reactions to Transport*

All the problems we studied required the simulation of chemical reactions, which occur on the membrane and in the bulk volume, combined with transport processes. We modeled this by using an operator splitting predictor/corrector method. A set of chemical reactions, specified for example in the Dashboard interface, was solved using the CVODE library for computing solutions to ordinary differential equations. This provided an input to a diffusive transport problem, which was solved by the methods described above. Iterating this process provided an algorithm for updating the time evolution of species concentration that locally balanced the effects of diffusion and transport.

We employed this algorithm to study the sporulation process in Caulobacter. The extreme stiffness of the differential equations in this case made the operator splitting technique difficult to apply. For this reason, or perhaps because the model itself contained many uncertainties our results did not confirm the conjectures of systems biologists nor did it validate our algorithms and software.

However, we also employed operator splitting to study the formation of lipid rafts as well as chemotaxis in neutrophils. In these cases we were able to validate the algorithm, affirm the results of earlier simulations, and show that the inclusion of realistic geometry had a significant effect on the operation of the model.

## Conclusions

In this work we have developed a number of key components for deterministic modeling of processes in cells, including diffusive transport coupled to chemical reactions, and an end-to-end process of grid generation from image data. The resulting methodology has a number of advantages over competing finite-element or unstructured-grid methods, including the ease of grid generation (which takes only a few minutes on a desktop workstation), and the efficiency and accuracy of the embedded boundary discretizations and solvers.

Even within the limited domain of reaction-diffusion models, there are still some outstanding technical questions that need to be addressed. Most prominent of them is the need for a replacement for operator splitting for the stiff coupling between surface reactions with diffusion in the bulk. Recent developments by Minion et. Al. in semi-implicit solvers based on the spectral deferred corrections approach of Dutt, Greengard, and Rokhlin show considerable promise as a possible solution. More fundamental issues include the modeling of cells with their full spatial and dynamic complexity, including mechanics, electrophysiology, and coupling to discrete processes such as microtubles and molecular motors. However, the approach presented here provides a solid foundation on which to build such extensions.